

# Detecção da COVID-19 Utilizando Imagens de Tomografia Computadorizada

Lucas Côgo Lampier

*Programa de Pós-Graduação em Engenharia Elétrica*  
*Universidade Federal do Espírito Santo*  
Vitória, Brasil  
lucas.lampier@edu.ufes.br

Leticia Araújo Silva

*Programa de Pós-Graduação em Engenharia Elétrica*  
*Universidade Federal do Espírito Santo*  
Vitória, Brasil  
leticia.silva.29@edu.ufes.br

Alberto Ferreira De Souza

*Programa de Pós-Graduação em Informática*  
*Universidade Federal do Espírito Santo*  
Vitória, Brasil  
albertodesouza@lcad.inf.ufes.br

**Abstract**—A doença coronavírus 2019 (COVID-19) apresenta questões em aberto sobre como diagnosticar e avaliar clinicamente o curso da doença. Recentemente, a tomografia computadorizada (TC) de tórax demonstrou utilidade para o diagnóstico de COVID-19. Neste trabalho, foram aplicadas três diferentes técnicas de deep learning em imagens de TC de voluntários diagnosticados ou não com COVID-19. Percebeu-se que a rede possuiu melhor desempenho quando a imagem de entrada continha as regiões de COVID-19 mais especificadas.

**Index Terms**—COVID-19, deep learning, tomografia

## I. INTRODUÇÃO

A pandemia do SARS-CoV-2 tem afetado milhões de pessoas em todo o mundo, resultando em mais de 170 milhões de casos confirmados e 3,7 milhões de mortes em 192 países ou regiões até o início de junho de 2021 [1], [2]. No Brasil, o número de casos confirmados é de, aproximadamente, 17 milhões, enquanto que o número de óbitos ultrapassa 470 mil [2].

A COVID-19 é uma doença infecciosa contagiosa que pode causar síndrome respiratória aguda grave, na qual a contaminação pode ocorrer devido ao contato com gotículas de saliva ou secreção nasal infectadas. Os sintomas mais comuns e leves causados pela COVID-19 são febre, tosse seca e cansaço. As mais graves são dificuldade de respirar ou falta de ar, dor ou pressão no peito e perda de fala e movimento [3]. Logo, o diagnóstico precoce da COVID-19 se torna uma tarefa urgente para prevenir a propagação da doença. A Reação da Transcriptase Reversa seguida pela Reação em Cadeia da Polimerase (RT-PCR) é um método amplamente utilizado para a detecção da COVID-19, em que os resultados podem ser afetados por baixa carga viral e erros de amostragem [1], [4].

Diversas pesquisas estão sendo realizadas a fim de buscar técnicas alternativas para diagnosticar a doença em pessoas afetadas [1], [4]–[6]. Neste contexto, equipamentos radiológicos, como raios-X e tomografia computadorizada (TC), surgiram como alternativas potenciais para o diagnóstico de COVID-19. No caso da TC, percebeu-se que ela pode ser

utilizada para diagnosticar a COVID-19, pois é mais sensível no diagnóstico de vírus do que o RT-PCR [1]. Os achados proeminentes da TC de tórax de pacientes de COVID-19 são opacidade em vidro fosco, consolidação multifocal irregular e um padrão de "pavimentação em mosaico" com uma distribuição periférica [7]. Em [1], um estudo foi realizado utilizando 351 imagens de pacientes diagnosticadas com COVID-19, e 396 diagnosticadas sem COVID-19. Eles utilizaram uma arquitetura de aprendizagem profunda denominada *Efficient-Net B4* e obtiveram uma acurácia (ACC) e *F1 score* de 90%. Uma ResNet-SVM foi utilizada em [4] utilizando 19 imagens de raio-X de tórax. Uma acurácia de 93% foi obtida, porém, o número de usuários da base de dados é muito pequeno. Em [5], uma rede neural convolucional (do inglês convolutional neural network, CNN) foi utilizada utilizando a imagem completa da TC do tórax. Eles obtiveram uma acurácia próxima de 80%. Um estudo foi realizado em [6] a fim de comparar 10 redes neurais convolucionais diferentes.

Desta forma, o objetivo deste trabalho é aplicar técnicas de Deep Learning em imagens de TC do tórax a fim de detectar a presença ou não da COVID-19 em pacientes.

## II. METODOLOGIA E RESULTADOS

### A. Descrição da Base de Dados

A base de dados consiste em 1110 TC scans divididos em 5 grupos: saudáveis (254 scans); com menos de 25% do pulmão infectado (684 scans); com infecção abrangendo 25% a 50% do pulmão (125 scans); com 50% a 75% (45 scans); e com mais de 75% do pulmão infectado (2 scans) [7]. Os dados foram obtidos utilizando o material disponibilizado no drive da disciplina de Deep Learning. Vale ressaltar que na base de dados utilizada, alguns dos scans estão faltando, porém, isto não interfere neste trabalho.

Apesar de estarem divididos em grupo saudável e grupos COVID-19, poucas são as amostras em que há marcações manuais da localização das infecções (máscaras). Nesta base, apenas 50 amostras são marcadas, sendo todas elas do grupo

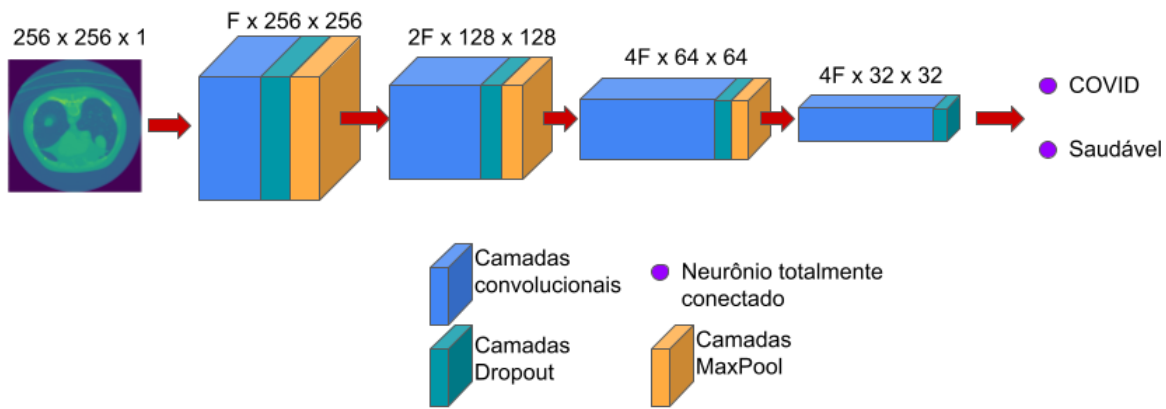


Fig. 1. Imagens completas.

de infecção com menos de 25% do pulmão infectado. Como esses são os frames com “certeza” de infecção da COVID-19 dentro dos scans de TC, eles foram utilizados. Ao todo, são 785 frames marcados com infecção, o que gera um dataset de 1570 frames. O que, a primeira vista, é um número razoável para a aplicação de deep learning.

Para gerar um conjunto padrão de treino e testes, os voluntários foram fixados como treino e teste em todos os testes. Desta forma, os voluntários utilizados para treinar os modelos do grupo COVID-19 são os voluntários identificados como 255 a 294, enquanto que os marcados como 295 a 304 foram utilizados para teste (todos estes são do grupo o com menos de 25% do pulmão infeccionado). Por fim, para o grupo dos voluntários saudáveis, os selecionados foram os com os códigos de 1 a 40 para treino, e de 41 a 50 para teste.

As TCs dos voluntários 255 a 304 possuem máscaras marcando a localização da infecção. Logo, os frames marcados são usados no grupo COVID-19 e os frames nas mesmas posições foram retirados do grupo saudável. No total, 1570 frames foram usados, sendo que 1252 para treino e 318 para teste, em que metade são frames contendo infecção por COVID-19 e metade de pulmões saudáveis. Todas as imagens foram normalizadas para valores entre 0 e 1 antes de serem utilizadas pelas redes.

### B. Teste 1: Imagens Completas

O primeiro teste consistiu em treinar uma rede baseada na arquitetura VGG para identificar os frames das TCs que capturaram infecção por COVID de frames de CTs capturados de pessoas saudáveis, utilizando a arquitetura base apresentada na Fig. 1.

Como dito, a arquitetura foi baseada na famosa VGG, com F sendo o número de filtros da primeira camada. Inicialmente foi testado com  $F = 64$ . As camadas dropout foram adicionadas como regularizadoras e para dificultar o overfitting. As camadas convolucionais foram construídas com filtros  $3 \times 3$ , e a função de ativação usada foi a Unidade Linear Retificada (do inglês Rectified Linear Unit, ReLU). As três primeiras camadas convolucionais são seguidas por uma camadas de max pooling  $2 \times 2$ , e uma “camada” dropout é

utilizada no treinamento com chance de 20% de zerar as saídas das camadas convolucionais. Os últimos dois neurônios são totalmente conectados com a função de ativação Softmax. O treinamento foi realizado no Google Colab utilizando a GPU fornecida de forma gratuita. A Rede foi treinada utilizando o otimizador Adam, e a métrica de perda calculada foi a entropia cruzada binária.

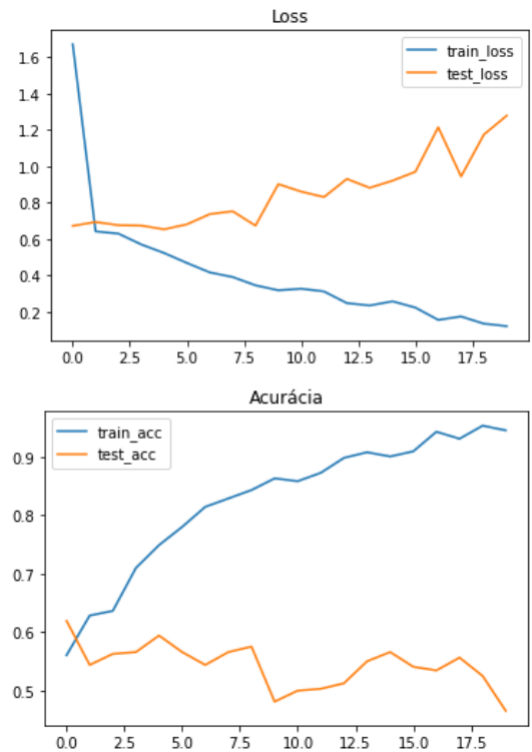


Fig. 2. Resultados utilizando as imagens completas.

A Fig. 2 mostra o resultado obtido no treino. Percebe-se que este não foi ideal, pois a perda de validação sobe logo nas primeiras épocas. A acurácia de validação é maior nas primeiras épocas, chegando a mais de 60% e depois caindo, e, conforme o modelo vai sofrendo overfitting, a acurácia

de validação tende a cair, enquanto a acurácia do conjunto de treino aumenta até chegar a quase 100%. Logo, isso indica que o modelo utiliza (ou não) outras características das imagens de treino além da infecção por COVID-19 para separar as imagens no treinamento. Uma possível abordagem para sobrepor este obstáculo seria diminuir a complexidade da rede para que as características relacionadas à infecção por COVID-19 possam se sobressair. O número de filtros das camadas convolucionais foram alterados para 8, 16, 32 e 32. O gráfico de treinamento do modelo é apresentado na Fig. 3. O otimizador utilizado foi o Adam com os parâmetros padrões do Keras (taxa de aprendizagem = 0.001,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 1e - 07$ ) e o cálculo de perda minimizado foi a entropia cruzada binária. Percebe-se que o resultado foi praticamente o mesmo. Então outra abordagem foi utilizada.

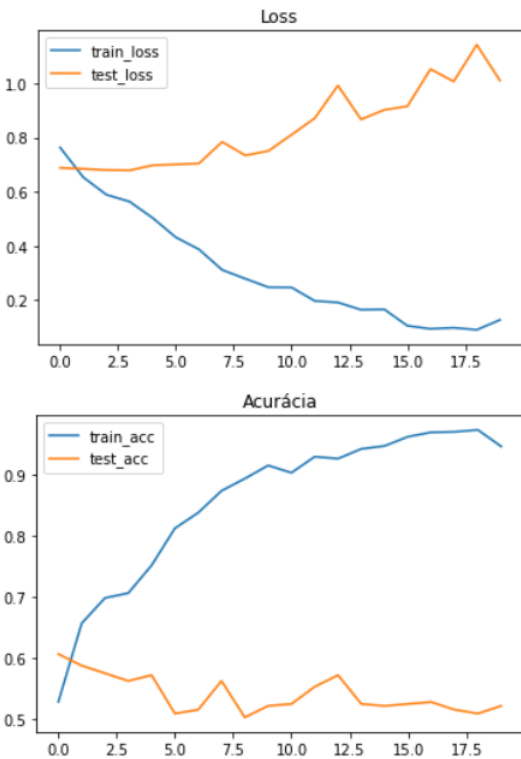


Fig. 3. Resultados utilizando as imagens completas simplificando a rede.

### C. Teste 2: Segmentar a Região dos Pulmões

Foi notado que o modelo sofreu overfitting rapidamente nas duas situações do teste anterior mesmo diminuindo a complexidade do modelo. Portanto, uma outra abordagem foi utilizada, que foi fazer a classificação usando dois modelos. O primeiro modelo foi realizado para selecionar a região do pulmão nas TCs, enquanto que o segundo modelo foi responsável por identificar a infecção por COVID-19 nos pulmões. Desta forma, como a região fora do pulmão é eliminada, diminui a quantidade de informações descorrelacionadas com a infecção que serão usadas no treinamento.

Para segmentar a região do pulmão, foi treinada uma U-net. A região do pulmão de 10 TCs (aproximadamente 400

imagens) de voluntários que não foram usados nos testes anteriores, foram segmentados para treinar e testar (8 para treino e 2 para teste) utilizando uma Deep Neural Network (DNN) para segmentação.

O modelo de U-net utilizado é apresentado na Fig. 4. Este modelo funciona como um “codificador e decodificador”: a primeira metade extrai informações importantes das imagens, e a segunda parte utiliza essas informações para reconstruir a imagem de entrada segmentando as partes desejadas. Neste modelo, a imagem passa primeiramente por três blocos contendo duas camadas convolucionais com função de ativação ReLU (no treinamento é adicionada uma camada de dropout de 20% entre elas) e uma operação de maxpooling 2 x 2. Em seguida, é realizada mais uma passagem convolucional. O processo é então “espelhado” fazendo blocos de camadas convolucionais e upsample (2x2), concatenando as saídas das operações de upsampling com as correspondentes das primeiras camadas de codificação. No final é adicionada uma camada convolucional com um filtro 1x1 para gerar a imagem de saída.

O gráfico de treinamento do modelo é apresentado na Fig. 5. O otimizador utilizado foi o Adam com os parâmetros padrão do Keras (taxa de aprendizagem = 0.001,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 1e - 07$ ) e o cálculo de perda minimizado foi a entropia cruzada binária.

É notável que o processo de treinamento do modelo de segmentação é bem lento. A perda de validação cai até a época 400 (aproximadamente 2h usando a placa gráfica do Google Colab).

Os resultados da segmentação foram promissores, por inspeção visual no grupo de testes, na grande maioria dos testes a região do pulmão foi bem segmentada do restante das imagens como exemplificado na Fig. 6, em que as máscaras geradas são multiplicadas pela imagem original. Em algumas poucas imagens, parte dos pulmões é segmentada, como se pode ver na Fig. 7.

As imagens usadas no primeiro teste (que são diferentes das usadas para treinar e testar a U-Net) foram então passadas na U-Net para segmentar a região dos pulmões e então a máscara gerada é multiplicada pela imagem original e os resultados são utilizados para treinar e testar as arquiteturas apresentadas no primeiro teste (com  $F = 64$  e  $F = 8$ ).

Os gráficos de treinamento do modelo são apresentados nas Figs. 8 e 9. O otimizador usado foi o Adam com os parâmetros padrão do Keras (taxa de aprendizagem = 0.001,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 1e - 07$ ) e o cálculo de perda minimizado foi a entropia cruzada binária.

Novamente a rede memoriza as características não relacionadas à COVID-19 para separar os dados de treino, novamente os resultados não são satisfatórios, em que as acurácias de treino chegam perto de 100% e as de validação ficam próximo dos 50%.

### D. Teste 3: Teste com as máscaras de Infecção por COVID

Foi percebido que nos testes anteriores os modelos atingem o overfitting rapidamente, mesmo em arquiteturas não muito

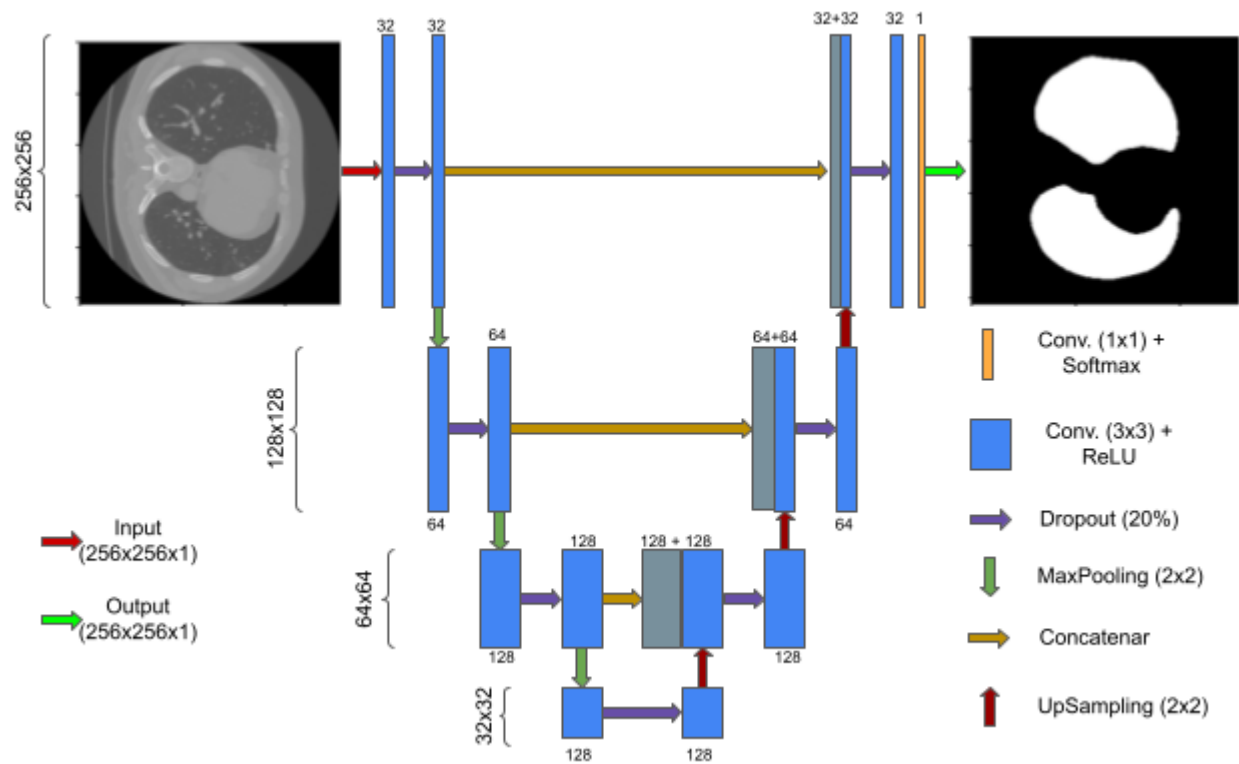


Fig. 4. Modelo utilizando U-Net.

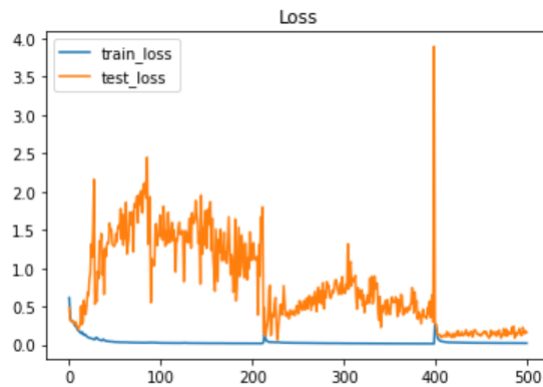


Fig. 5. Resultados obtidos no Teste 2.

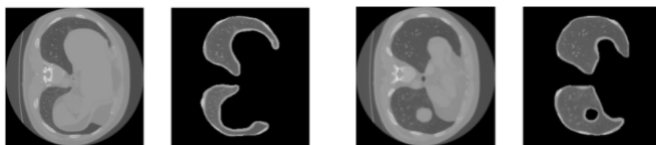


Fig. 6. Exemplos de imagens segmentadas.

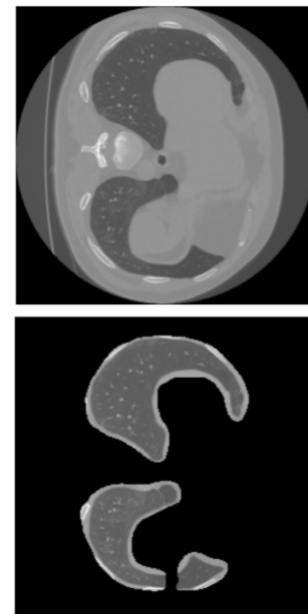


Fig. 7. Exemplo de imagem com parte do pulmão segmentado.

complexas, enquanto que no grupo de validação a acurácia dificilmente ultrapassa os 60% nas melhores épocas. Percebe-se que há muita informação, além da infecção por COVID-19, sendo necessário uma base maior para balancear as informações e os dados de COVID-19. Então, uma nova

abordagem foi realizada de forma a indicar para o modelo a parte da TC que deveria ser avaliada para identificar os que contém a infecção. Então, neste terceiro teste, as máscaras de 50 TCs fornecidas pelo banco de dados são utilizadas em uma nova U-Net - que desta vez será treinada para identificar

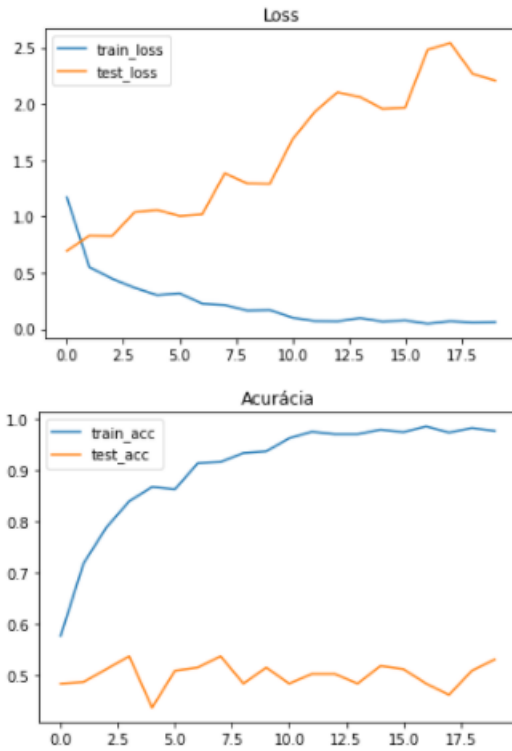


Fig. 8. Gráficos de acurácia e de perda do segundo teste utilizando a arquitetura VGG com 64, 128, 256 e 256 filtros nas camadas convolucionais.

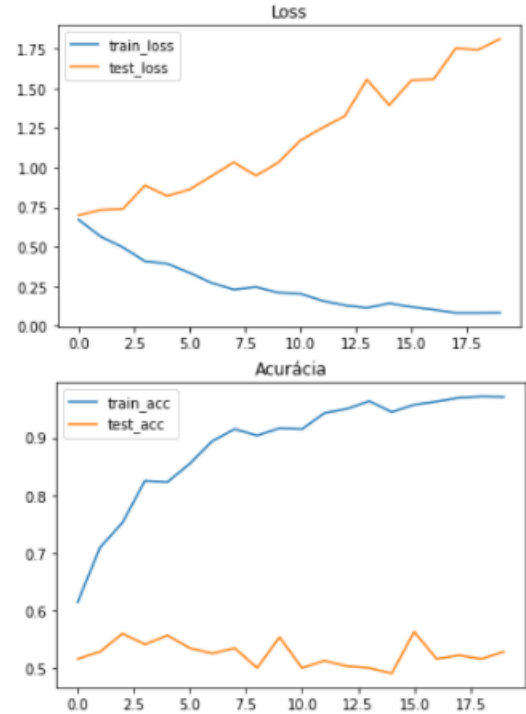


Fig. 9. Gráficos de acurácia e de perda do segundo teste utilizando a arquitetura VGG com 8, 116, 32 e 32 filtros nas camadas convolucionais.

diretamente as regiões infectadas.

Em cada TC existe aproximadamente 40 frames, na quais a infecção aparece em aproximadamente 40% deles. Então, das 50 TCs com marcações, a U-Net será treinada para identificar especificamente a região infectada apenas quando esta existir. O modelo de U-Net é o mesmo apresentado no Teste 2, entretanto, para tornar o treinamento mais rápido, as imagens foram reduzidas de 512 x 512 para 128 x 128 (ao invés dos 256 x 256 utilizados anteriormente).

Parte das máscaras é apresentada na Fig. 10.

Os gráficos de treinamento do modelo são apresentados na Fig. 11. 40 TCs (aproximadamente 1600 frames) foram utilizadas para treino e 10 TCs (aproximadamente 400 frames) para teste. O otimizador utilizado foi o Adam com os parâmetros padrão do keras (taxa de aprendizagem = 0.001,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 1e - 07$ ) e o cálculo de perda minimizado foi a entropia cruzada binária.

O gráfico de acurácia não revela muita coisa neste treinamento já que a grande maioria dos pixels do TC é de pixels não-COVID-19. No gráfico da perda, é notável que a perda do grupo de testes foi atenuada até a época 70, aproximadamente, e depois se estabilizou.

Para avaliar a eficácia do modelo (comparando-o aos outros testes), a acurácia foi medida da seguinte forma: caso o modelo tenha encontrado pelo menos um pixel no frame que foi classificado como pixel “Covid”, todo o frame é categorizado no grupo COVID-19 (grupo 1); caso contrário o frame é

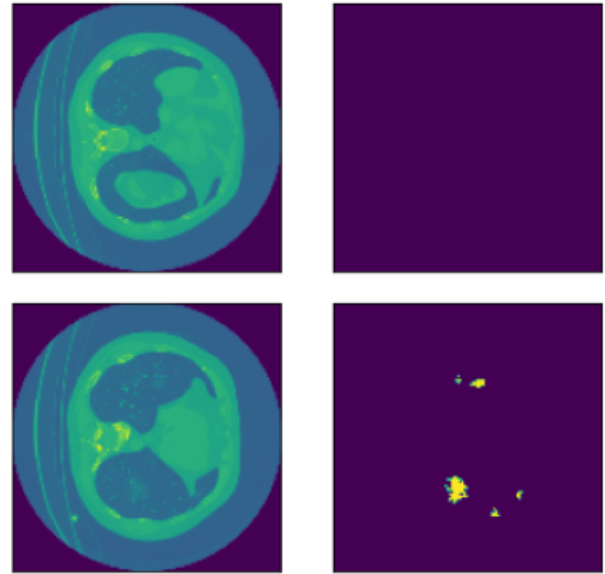


Fig. 10. Exemplos de máscaras.

categorizado no grupo saudável (grupo 0). Utilizando essa métrica, a rede categorizou corretamente 82% dos frames no grupo de teste (bem superior aos testes testando separar diretamente os frames COVID-19 e saudáveis). Para uma detalhamento melhor a matriz de confusão é apresentada na Tabela I.

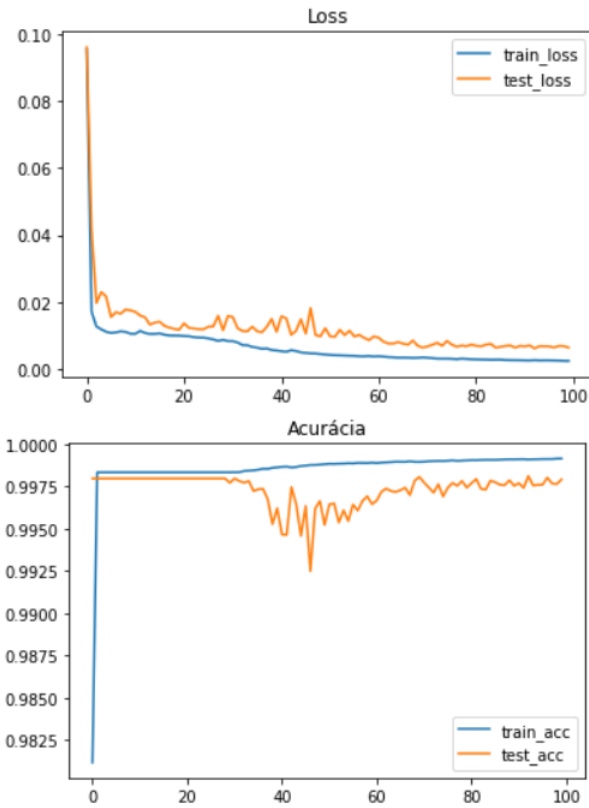


Fig. 11. Resultados obtidos no Teste 3.

TABLE I  
MATRIZ DE CONFUSÃO

	Saudável	COVID-19
Saudável	216	35
COVID-19	37	121

Pela tabela temos as seguintes métricas do modelo (Saudável = falso, COVID-19 = verdadeiro):

$$Acurácia = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Sensitividade = \frac{TP}{TP + FN} \quad (2)$$

$$Especificidade = \frac{TN}{TN + FP} \quad (3)$$

Onde, TP são os verdadeiros positivos, TN os verdadeiros negativos, FP os falsos positivos, e FN os falsos negativos. A Tabela II resume os resultados obtidos utilizando as métricas citadas.

### III. CONCLUSÃO

Esse trabalho apresentou uma proposta de aplicação de técnicas de deep learning no processamento de imagens de TC de pessoas diagnosticadas com COVID-19. Após a análise dos resultados, percebeu-se que a rede aprende melhor quando foi

TABLE II  
RESULTADOS OBTIDOS.

Acurácia	82.40%
Sensitividade	76.58%
Especificidade	86.06%

dado a ela da maneira mais específica o possível os pontos das imagens que ela deve usar para identificar as regiões infectadas por COVID-19.

### REFERENCES

- [1] T. Anwar and S. Zakir, "Deep learning based diagnosis of COVID-19 using chest CT-scan images," 2020 IEEE 23rd International Multitopic Conference (INMIC), 2020, pp. 1-5, doi: 10.1109/INMIC50486.2020.9318212.
- [2] "WHO Coronavirus (COVID-19) Dashboard," World Health Organization. [Online]. Available: <https://covid19.who.int/>. [Accessed: 08-Jun-2021].
- [3] "Coronavirus," World Health Organization. [Online]. Available: [https://www.who.int/health-topics/coronavirus#tab=tab\\_1](https://www.who.int/health-topics/coronavirus#tab=tab_1). [Accessed: 08-Jun-2021].
- [4] C. Zhou, J. Song, S. Zhou, Z. Zhang and J. Xing, "COVID-19 Detection based on Image Regrouping and ResNet-SVM using Chest X-ray Images," in IEEE Access, doi: 10.1109/ACCESS.2021.3086229.
- [5] E. H. Lee *et al.*, "Deep COVID DeteCT: an international experience on COVID-19 lung detection and prognosis using chest CT," npj Digital Medicine, vol. 4, no. 1, 2021.
- [6] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks," Computers in Biology and Medicine, vol. 121, p. 103795, 2020.
- [7] S. P. Morozov *et al.*, "MosMedData: Chest CT Scans with COVID-19 Related Findings Dataset," medRxiv, p. 2020.05.20.20100362, Jan. 2020.